

AD-A251 598



DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

(2)

ion is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this burdening this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

2. REPORT DATE July 16, 1991		3. REPORT TYPE AND DATES COVERED FINAL 1 Jun 89 - 31 May 91	
4. TITLE AND SUBTITLE "Design & Training of Limited-Interconnect Architectures" (U)		5. FUNDING NUMBERS 61102F 2305/B3	
6. AUTHOR(S) Lex A. Akers Mark R. Walker Siamack Haghighi		AFOSR-TR- 92 0004	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Arizona State University College of Engineering/Applied Sciences Elec. Engineering, Center for Solid State Elec. Research Tempe, AZ 85287-6206		8. PERFORMING ORGANIZATION REPORT NUMBER DWA 1752	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Air Force Office of Scientific Research Building 410 Bolling AFB, DC 20332-6448 nm		10. SPONSORING/MONITORING AGENCY REPORT NUMBER AFOSR-89-0430	
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; Distribution unlimited		12b. DISTRIBUTION CODE UL	
13. ABSTRACT (Maximum 200 words) This DRAFT Final Report describes the progress on the AFOSR research contract on the design and training of limited-interconnect neural architectures. A novel local training rule has been derived, an information theoretic analysis to determine capabilities and limitations of multi-layered limited-interconnect neural networks was developed, and the hardware implementation of local training rules analyzed.			
14. SUBJECT TERMS Neural Networks, Information Theory		15. NUMBER OF PAGES 5	
		16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

Final Report

Design and Training of Limited-Interconnect Architectures

Lex A. Akers
Center for Solid State Electronics Research
Arizona State University

Abstract

This report describes the progress on the AFOSR research contract on the design and training of limited-interconnect neural architectures. A novel local training rule has been derived, an information theoretic analysis to determine capabilities and limitations of multi-layered limited-interconnect neural networks was developed, and the hardware implementation of local training rules analyzed.

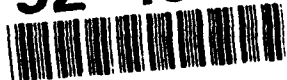
1. Introduction

In recent years, there has been a resurgence of interest on neural networks and models. The potential of massive parallelism of such networks have generated this interest.. Neural networks are best suited for such applications as pattern recognition and signal processing. Neuromorphic (brain like) models, allow an alternative for achieving real-time operation for such tasks, while having a compact and robust architecture.

Neuromorphic models consist of interconnections of simple computational nodes. In this approach, each node computes a weighted spatio-temporal integration of its inputs, and makes a decision based on this integration. Specification of a neural network model will consist of two parts. The first part is the architecture specification. This includes determining the number of nodes and the interconnection pattern of these nodes. The second specification task is providing a training rule. The purpose of the training rule is to allow a network of nodes to perform in a desired collective manner. This approach offers many advantages such as:

1. Since the nodes in a neural network may operate locally on the incoming information, a high degree of parallelism is achieved. The requirement of an adaptation rule, for a network of interconnected neurons, is one of the many advantages of neural networks over existing architectures.

92 6 15 154

92-15625


2. Due to distributed representation of information in neural networks, failure of a node or a connection will not cause severe degradation of the network performance. Therefore, a highly robust system can be obtained.
3. Since the computations are done in parallel, real-time hardware operational performance can be achieved.

At present, most multi-layered neural network models assume full inter-layer connectivity. This requirement severely limits the maximum size of VLSI or system implementation of such architectures. Moreover, such issues as the limited resolution of weights and local computations, will make many neural models unattractive for hardware implementation. Finally, scaling property for larger size networks, need to be taken into consideration.

In an effort to overcome the above difficulties, our research efforts have been focused on neural networks that can be implemented in VLSI hardware. This is especially important for real-time operational performance.

II. Research Objectives

The research objectives were:

1. Development of on-chip local training rules specifically designed for optimal operation in hardware,
2. Use information theoretic analysis to determine capabilities and limitations of multi-layered limited-interconnect neural networks, and
3. Analysis of hardware implementation issues such as finite precision of weight storage and multiplications, compact cell development, scaling of architectures to large numbers of neurons, and fault tolerance.

A. Information-Theoretic Analysis of Limited-Interconnect, Finite Parameter Neural Networks

Information theory provides a formalism for the analysis of network architecture and adaptation. It can be demonstrated that a unifying organizing principal of network internal representations is maximum entropy. For example, processing for improved linear separability, increased data orthogonality, projection of data onto higher dimensional space, are all specific instances of increased entropy data representations.

Increased entropy can be shown to be an artifact of parameter constraint - eg. round off will increase entropy. An optimal hidden unit representation (achieved by increasing the entropy of the input) can be accomplished independent of the network implementation parameters, such as connectivity and resolution.

Traditionally, neural networks are used as "black boxes" where optimal network hidden unit size is determined experimentally. By employing entropy as a qualitative measure of hidden layer performance, an optimal network architecture may be obtained by "stacking" elements until entropy is maximum for the given training set. This may permit substantial resource



Availability Codes	
Dist	Avail and/or Special
A-1	

savings: by employing limited fan-in nodes, an optimal hidden unit representation (the layer of maximum entropy) may be discovered below the point of full connectivity (each node in a given layer receives a signal for all input bits).

B. Reasons for Local Unsupervised Adaptation Rules

Local connectivity on the input field is found commonly in the biological receptive field allowing a single trend, mode, frequency, feature, etc. to represent a subset of the input. Supervised adaptation requires global routing of error signals consuming large areas of the semiconductor chip.

C. Advantages of Local Unsupervised Adaptation Rules

A local processor's operation can be optimized with unsupervised adaptation. Theoretical analysis and designs are facilitated when low fan-in modular processing elements can be used. Neuromorphic computation emerges when local processors are connected.

III. Accomplishments

- We have derived a new local unsupervised training rule based on the optimization of the mutual information among the inputs to a computing node.
- A network of neurons with weights trained using this local training rule allows the integration of local information to arrive at the global representation of the information.
- We have also shown that for various arrangements of the weights, optimal response to certain classes of patterns can be obtained.
- The neural network developed with this on-chip training rule requires only local connections between neurons, is highly parallel in nature, and allows continuous and real-time adaptive VLSI implementation.
- The advantage of this model over other existing models is that a multitude of behavior such as edge detection, image compression, and dynamic object detection can all be achieved with the same architecture in continuous time. Other advantages are the compact size, and robustness to random noise.
- We have also developed a real-time VLSI smart sensor. The device consists of a low-resolution optical sensor array, and a high-resolution optical feature detector.
- A smart-scan mechanism is employed, in which the low-resolution sensor array examines regions of the object's image for the presence of corners. Once it detects the presence of a corner, a high-resolution feature detector which is analogous to the high-resolution fovea of the retina, scans the region.
- The high-resolution detector develops an encoding of the image by sensing and storing all corners within the object and their respective orientations. The orientations and angles of the corners in an object may be employed to generate a nearly unique abstraction of the object. This scheme is particularly useful as corners are spatially-localized and hence their detection

requires only local information.

- The classification of the analog vector generated from the corner detecting process is accomplished using the k-means clustering technique. A database of clusters is maintained and updated as required.
- An information theoretic formalism developed allows us to analyze neuromorphic processing occurring in different neural architectures.
- The rate of information refinement as a function of local fan-in, resolution of the components, and number of neural processing layers has been developed.

IV. Publications

A. Proceedings and journal papers

1. M. Walker, S. Haghighi, A. Afghan, and L. Akers, "Training of a Limited-Interconnect, Synthetic Neural IC", Neural Information Processing Systems 1, Editor. E. Touretsky, Morgan Kaufmann Pub., 1989.
2. M. Walker, P. Hasler, and L. Akers, "An Adaptive, CMOS Neural Array for Pattern Association", Proceeding of the Phoenix Conference on Computers and Communications April 1989.
3. L. Akers, M. Walker, R. Grondin, and D. Ferry, "A Synthetic Neural Integrated Circuit", Proceeding of the Custom Integrated Circuits Conference, May, 1989.
4. L. Akers, M. Walker, S. Haghighi, R. Grondin, and D. Ferry, "Parallel Hardware Architecture for Neuromorphic Computation", Proceeding of the Third Annual Parallel Processing Symposium, March 1989(invited paper).
5. A. Rao, and L. Akers, "ART1 Network Implementation Issues", Proceeding of the TENCON 89 Conference, 1989.
6. A. Rao, M. Walker, and L. Clark, L. Akers, "Integrated Circuit Emulation of ART1 Networks", Proceeding of the IEE, 1989.
7. M. Walker, P. Hasler, and L. Akers, "CMOS Neural Networks for Pattern Association", IEEE Micro, Oct. 1989(invited paper).
8. A. Rao, M. Walker, L. Clark, L. Akers, and R. Grondin, "VLSI Implementation of Neural Classifiers", Neural Computation, 2, p.35-43, 1990.
9. A. Rao, S. Haghighi, and L. Akers, "Electronic Implementation of Synthetic Neural Systems", Artificial Life Conference Program Abstracts, Santa Fe, NM, Feb. 1990.
10. P. Hasler, and L. Akers, "VLSI Neural Circuits and Systems", Proceeding of the Phoenix Conference on Computers and Communications April 1990.
11. A. Rao, and L. Akers, "Retinomorphic VLSI Smart Sensor for Invariant Geometric Object

Recognition", Proceedings of the IEEE International Conference on Neural Networks, June 1990.

12. S. Haghighi, and L. Akers, "TRELIS: An Optimal Unsupervised Local Training Rule", Proceedings of the IEEE International Conference on Neural Networks, June 1990.

13. P. Hasler, and L. Akers, "VLSI Neural Systems and Circuits", International Journal of Computer Aided VLSI Design, (in press).

14. A. Rao and L. Akers, "A VLSI-Implementable Scheme for Invariant Object Recognition", Proceeding of the Phoenix Conference on Computers and Communications, p.17, April 1991.

15. P. Hasler, and L. Akers, "Implementation of Analog Neural Networks", Proceeding of the Phoenix Conference on Computers and Communications, p.32, April 1991.

16. M. Walker, and L. Akers, "Information Refinement in Unsupervised, Implementation Constrained Neural Networks", Proceeding of the AMSE International Conference on Neural Networks, May 1991 (with M. Walker).

17. M. Haghighi, and L. Akers, "An unsupervised training rule for dynamic information processing", Proceedings of the IEEE International Conference on Neural Networks, July 1991..

18. "An Continuous time synapse employing a refreshable multilevel memory ", Proceedings of the IEEE International Conference on Neural Networks, July 1991..

B. Book Chapters

1. "Two Dimensional Automata in VLSI", Integrated Circuits in .5 to .05 Micron Dimensional Range, R.K. Watts, Ed., John Wiley, New York, 1990 (D. Ferry, L. Akers, and R. Grondin contributing authors).

2. "VLSI Implementation of Neural Structures", The Computer and the Brain, J. Brink and D. Nagy, Eds. MIT Press, 1990 (L. Akers, D. Ferry, and R. Grondin contributing authors).

3. "A Limited-Interconnect, Highly Layered Synthetic Neural Architecture", VLSI for Artificial Intelligence, Kluwer Academic Press, 1989. L. Akers, M. Walker, D. Ferry, and R. Grondin contributing authors).

4. "Synthetic Neural Systems in VLSI", An Introduction to Neural and Electronic Networks, Academic Press, 1990 (L. Akers, D. Ferry., and R. Grondin contributing authors).

5. "Architectures and Devices for ULSI" , Frontiers in Computing Systems Research, Plenum 1990 (D. Ferry, R. Grondin, and L. Akers contributing authors).

6. "VLSI Implementations of Sensory Processing Systems", Advanced Neural Computers, R. Eckmiller, Eds. Elsevier Science Publishers, New York 1990 (S. Highighi, A. Rao and L. Akers contributing authors).